

Stroke ICU Patient Mortality Day Prediction

Oleg Metsker¹[0000-0002-6231-8036], Voznjok Igor², Georgy Kopanitsa¹[0000-0002-6231-8036],
Elena Morozova² and Prohorova Maria²

¹ ITMO University, Saint Petersburg, Russia

² The Saint Petersburg Research Institute of Emergency Medicine. I.I. Dzhanlidze
olegmetsker@gmail.com, voznjouk@yandex.ru,
georgy.kopanitsa@gmail.com, novaj44@mail.ru, airty@mail.ru

Abstract. This article presents a study on development of methods for analysis of data reflecting the process of treatment of stroke inpatients to predict clinical outcomes at the emergency care unit. The aim of this work is to develop models for the creation of validated risk scales for early intravenous stroke with minimum number of parameters with maximum prognostic accuracy and possibility to calculate the time of "expected intravenous stroke mortality". The study of experience in the development and use of medical information systems allows us to state the insufficient ability of existing models for adequate data analysis, weak formalization and lack of system approach in the collection of diagnostic data, insufficient personalization of diagnostic data on the factors determining early intravenous stroke mortality.

In our study we divided patients into 3 subgroups according to the time of death - up to 1 day, 1 to 3 days, and 4 to 10 days. Early mortality in each subgroup was associated with a number of demographic, clinical, and instrumental-laboratory characteristics based on the interpretation of the results of calculating the significance of predictors of binary classification models by machine learning methods from the Scikit-Learn library. The target classes in training were "mortality rate of 1 day", "mortality rate of 1-3 days", "mortality rate from 4 days". AUC ROC of trained models reached 91% for the method of random forest. The results of interpretation of decision trees and calculation of significance of predictors of built-in methods of random forest coincide that can prove to correctness of calculations.

Keywords: ICU, Stroke, Mortality, Machine learning, Mortality prediction.

1 Introduction

Stroke is the second most deadly cause of death worldwide. In Russia, brain stroke is the second leading cause of death after myocardial infarction. Every year around 450000 people suffer from stroke, in fact it is the population of a big city [1]. The mortality rate in Russia is 4 times higher than in the USA and Canada [2]. Among European countries, the mortality rate from cerebrovascular diseases is the highest in Russia. According to the All-Russian Center for Preventive Medicine, in our country 25% of men and 39% of women die from cerebrovascular diseases. In the largest cities

of the country the situation with this type of pathology is extremely unfavorable. In St. Petersburg, for example, the frequency of stroke is about 528 cases per 100,000 residents, while the mortality rate for ischemic stroke is 39%. It is necessary to emphasize the catastrophic consequences of ischemic stroke - up to 84-87% of patients die or remain disabled and only 16-13% of patients fully recover [2]. According to the findings of a large-scale study of recent years, some modern epidemiological trends have been identified [3]: In general, global statistics show a decline in stroke mortality over the past two decades due to the introduction of new treatments (thrombolysis, thrombectomy), but the absolute number of people who have stroke is only increasing every year [4]. This nosology is still the strong second leading cause of death from cardiovascular disease (CVD), remaining the undisputed leader among all nosologies leading to severe disability. Hospital mortality remains one of the most important quality indicators, which can be used to identify problems associated with the optimization of prehospital and hospital treatment processes. It can be used to assess the effectiveness of primary and secondary care, routing, and the degree of implementation of modern diagnostic and treatment algorithms, including the quality of interaction between different levels of care [5]. It is important to note that regional characteristics of the populations may significantly differ from the global ones, and the development of specialized care programs for patients with a stroke has its national and institutional characteristics. Understanding the factors that contribute to the reduction of hospital mortality will allow us to develop a targeted strategy for the development of services providing care to patients with a stroke in Russia and in the world. Thus, development of personalized models and algorithms for planning of individual treatment tactics for the stroke patients can reduce mortality and increase the standard of life. The development of such models and algorithms will ensure better continuity and efficiency of medical care and help reducing the number of complications. The basis for such models can be the scales of calculation of patients' mortality risks in emergency units, which are also absent in Russia at present.

2 Related works

Most statistics are accumulated in national stroke registries or national databases: China National Stroke Registry II (CNSR II) [6], the Nationwide Hospital Discharge Database (NHDD), Berlin Stroke Register (BSR), German Stroke Register, the Registry of the Canadian Stroke Network (RCSN), National Acute Stroke Israeli (NASIS) registry, FLENI Stroke Data Bank, Australian Stroke Clinical Registry (AuSCR), National Stroke Register of Ireland, the Austrian Stroke Registry. The analysis of available literature revealed rather heterogeneous values of the share of hospital mortality of patients with stroke in different countries. At the same time, direct indicators of the share of hospital mortality had significant differences from 1,4% in China [7] to 22,7% in Ethiopia [8]. Significant differences in data can be explained both by the quality of care and by the nature of statistical data collection. In particular, most of the reports took into account only the ischemic type of stroke [9–14], different exclusion criteria were applied in a number of observations - daily mortality and stay exceeding 180 days [9],

in-hospital stroke [15], patients in need of admission to the general intensive care unit [16] or a general department. It should also be noted that samples are heterogeneous in terms of the number of patients: from 110 [8] to 12 million patients [17]. Hospital mortality rates vary considerably between facilities within the same country. For example, the average hospital mortality rate in Germany in 2011 was 4.6% when 26 Stroke units were evaluated. [13], at the same time as in the German study of 2015 on this parameter was 8.2% [18]. In Australia, hospital mortality also varies significantly (from 7% to 23%) depending on the level of the hospital. [15], in Germany, there is a dependence on the size of the hospital - from 0% to 25% in small hospitals and from 0.4% to 9.3% in large hospitals. [11]. Only 9 studies out of 22 provide data that allow tracking the dynamics of changes in the indicator of intra-hospital mortality. The average rate of decline in this indicator was 0.36% per year. Rapid changes in this parameter are more typical of the ischemic type of stroke, and mainly the faster rate of decline was associated with the introduction and expansion of the vascular center network for stroke (with mandatory Stroke Unit). The most significant example of Canada - where vascular center system was introduced, which led to the rate of change in the provinces was 0.28% per year, while in the provinces without the introduction of the vascular center system, the rate changed only by 0.11% per year [19]. The availability of prognostic models and scales that are understandable to clinical staff and easy to operate, reduces hospital mortality and allows for a more targeted and individualized approach to therapy. Such models should take into account locally established practices. Models should be available that can predict a fatal scenario for the disease, considering all relevant factors. To date, the international medical community has made repeated attempts to create such a prognostic scale. In the 2002 review, C. Counsell and M. Dennis analyzed 83 models with a total of 150 prognostic factors, and the assessment resulted in only 4 models meeting quality criteria [8]. The databases have a huge number of parameters including various tests and indicators. In some cases, the use of a large number of features leads to lower rates of learning and forecasting, reduces the predictive accuracy of the model, and prevents the model from being interpreted, which is an important requirement for models used in medicine. Thus, finding the best set of features in the context of our task is one of the key factors ensuring high quality of the predictive model.

On the basis of the analysis of 12 modern prognostic models from 10 countries we can identify some of the most stable (main) predictors for the causes of intra-hospital mortality: age [16,20–24]; type of stroke [25]; lesion location [25]; level of consciousness [11,20,23,25,26] upon admission; NIHSS stroke severity [10,21,22,24]; comorbidity [22,27], Charlson comorbidity index [23], Atrial fibrillation [11][22], case history Transitor ischemic attack (TIA) [31]; hospital complications (high intracranial pressure) [16], pneumonia, seizures, anxiety/depression, infections, limb pains and constipation [22,27].

Among the predictors related to the organization of care, the time of admission to hospital can be noted - in a Japanese study, the 7-day mortality rate increased if the patient was admitted on weekends or holidays. [23], hospital delivery method had a predictive value as well [16], Both these parameters are included in the GWTG-Stroke program [14]. In order to identify priority areas for improving the outcome of the dis-

ease it is necessary to divide the selected factors (predictors) into modifiable and unmodifiable, respectively. Modifiable mortality predictors can be referred to: time and method of hospital delivery; qualifications of medical personnel; stroke care model; history of stroke or TIA, atrial fibrillation, diabetes mellitus, comorbidity index - parameters to which primary prevention should be directed; intra-hospital complications (high intracranial pressure pneumonia, seizures, anxiety/ depression, infection, extremity pain and constipation). A special form of complications in the form of extracerebral pathology - polyorgan failure syndrome - is distinguished separately. Special attention should be paid to the prevention of this syndrome. The unmodifiable factors of stroke mortality include: gender, age, type of stroke, localization of lesion. As for the assessment of the impact of comorbid diseases, it is important to consider not only the presence of individual pathologies, but also their combination. In particular, the following groups can be distinguished: Arterial hypertension + atrial fibrillation, Arterial hypertension + atrial fibrillation + Coronary heart disease, atrial fibrillation + postinfarction cardiosclerosis, and, Arterial hypertension + postinfarction cardiosclerosis + atrial fibrillation MA и + Diabetes mellitus. Only two studies presented clear prognostic scales containing a scoring system for rapid assessment of the risk (probability) of in hospital mortality [14,21]. The PREMISE scale is simple, quick to calculate at > 85% of strokes and uses only variables that are available shortly after the onset of ischemic stroke when admitted to the Stroke Unit. It should be noted that the practical application of any analyzed scale above in different countries requires corrections to take into account regional peculiarities - social, geographical and medical and economic factors. [26]. The creation of such scales and models in Russia would provide a tool to assess the efficiency of care. The goal of this work is to identify features for the creation of validated risk scales for early hospital mortality.

3 Methods

3.1 Cohort description

The study includes data about 36450 episodes (17158 outpatient 5565-inpatient patients 200-lethal patients 5565 patients who has international criteria for diagnosis ICD i60 to i69.8) and were treated in the Almazov national research center from 2011 to 2019. Among the causes of admission: ischemic stroke, hemorrhagic stroke, embolic stroke, transitor attacks. As the initial data describing the condition the patient's examination data at the intake and use of clinical scales (NIHSS, mRS), conclusion of magnetic resonance imaging (MRI), conclusion of ultrasound investigation, data from laboratory tests, data on treatment events from the medical information system. A separate more detailed analysis of the group of only deceased patients from 100 people was carried out to identify differences and mortality factors in different time periods (1 day, 2-3 days, 4-15 days) on the basis of data from the The Saint Petersburg Research Institute of Emergency Medicine n.a. I.I. Dzhanlidze¹.

¹ <http://www.emergency.spb.ru/>

The data of the medical information system of the operating specialized center of MRI, ultrasound and other characteristics of the volume of cerebral and vascular stroke examination were compared with the data on the duration and outcomes and time of death. Information about hospital mortality was included in the study, if they met the following criteria: the fact of clinically confirmed diagnosis of acute cerebral circulation disorder (ischemic or hemorrhagic), with the presence of focal, general cerebral neurological syndromes, which lasted more than 24 hours from the beginning of the disease; hospitalization in connection with stroke in the first day of the disease; the entire period of hospitalization in connection with acute case of the patient spent in one institution; lethal outcome was associated with an acute period of stroke. Information confirming lesions of the brain substance has been obtained from data from the CT scan and/or MRI of the brain, which have been repeated if necessary. The extent of precerebral and cerebral artery lesions was assessed using ultrasound duplex scanning, CT scan, MRI or cerebral angiography.

3.2 Analysis and machine learning methods

To obtain the optimal set of features a combination of classical methods based on different correlation coefficients of features (Pearson correlation coefficient and Spearman correlation coefficient) were used. Ensemble algorithms, including ensembles built on the basis of models with the use of decision trees, and random forest are used as prognostic models. A Scikit-Learn library was used to implement machine learning methods. In the process of definition of hyperparameters of machine learning models, cross-validation by k blocks was applied. Precision and Recall (accuracy and completeness), as well as their harmonic mean (F-score) were used as metrics at this stage. Construction of the confusion matrix of multiclass classification allowed to analyze errors, improve data sampling used for model training and initialize the next iteration of model training. The data on treatment of real patients from the Almazov Center were used for validation of the final resulting models. The data of patients who did not participate in any stages of model training and adjustment of hyper-parameters were used. AUC ROC - the area under the error curve - was used as the result metrics. P-value was calculated using two methods. The essence of the first method is that for each sample of dead (<1 day, 1-3 days, 4-10 days) we have calculated P-value for every feature of the corresponding test. Chi-square criterion was used for categorical features, Kolmogorov-Smirnov's test was used for continuous features. The essence of the second method of calculating P-value by one attribute (mortality period) for three groups of patients according to the severity and type of stroke (group 1: ICH+PVH, IS+ICH; group 2: IS+Bilat atr, Is-foc 16-25hu; group 3: IS-1/3<16HU, Sub Tent ICV).

4 Results

The analysis obtained a general model of mortality for all patients with stroke AUC ROC-93% demonstrated random forest learned on the dataset with more than 60 laboratory and personal patient observation features. Three separate models have also been

developed for patients with different lethality periods (up to 1 day, from 1 to 3 days, and from 4 to 15 days) using decision trees that showed an AUC ROC of 85 to 91 %. For these models, the dataset consisted of more than 70 specialized features, including a score on neurologic scales, brain examination data, assessment of the patient's consciousness and somatic state. The importance of features for different duration of lethality was also compared. Moreover, a clinical interpretation of the comparison results is given below.

4.1 General machine learning model

The models were trained on the dataset describing 5565 patients who were treated as a binary classification models by machine learning methods from the Scikit-Learn library. The following parameters were used as features: patients age, male, pressure, area of brain damage, the size of the hematoma. Moreover, the following laboratory tests were used as features: MCHC-red blood cell index, endothelin, interleukin-10, interleukin-8, interleukin-6, interleukin-4, interleukin-1-beta, INR, fibrogen, vitamin D, parathormone, urine Nitrites, urine bilirubin, urine, bld urine, LEU urine, urine NIT, urine KET, urine glucose, urine PRO, urine Ph, urine color, D-dimmer, albumin, lipids, triglyceride, total cholesterol, prothrombin index, fibrinogen by Klaus, K+ (Vienna), neutrophils, monocytes, lymphocytes, MPV average, platelet volume, PDW Width of platelet distribution by volume, RDW Width of red blood cell distribution by volume, MCHC the average concentration of hemoglobin in eritr, MCH is the average hemoglobin content in 1 erythrocyte average volume of red blood cells, reactive protein, erythrocyte sedimentation rate, troponin, ALT, AST, HGB Hemoglobin, WBC white blood cells, RBC red blood cells, PLT platelets, creatinine, bilirubin, HCT Hematocrit, glucose level.

Random forest demonstrated the best AUC ROC-93%. The nine most importantly lethality features of the stroke patient further: systolic pressure (0.06), RBC red blood cells (0.05), interleukin-8 (0.05), HCT hematocrit (0.04), diastolic pressure (0.04), age (0.03), MCHC - red blood cell index(0.03), ventricular damage(0.03), hematoma volume (0.03).

The following conclusions emerge from the general analysis of the overall data: 1. *Terms of mortality.* All cases of death of patients, which were distributed within 14 days, were estimated, with the greatest number of lethal outcomes occurring within 5 days. 2. *Patients age* 60 to 90 years (at least 65% out of 200 dead) were most frequently encountered in the group of the deceased, the maximum frequency (25%) falls on the age of 60 to 70 years, in the same age group there is the maximum morbidity of stroke with their share is almost 35% of the number of diseased. 3. *Among the deceased,* men prevailed (by more than 25%). 4. *The proportion of deaths* in the hemorrhagic stroke cohort was twice as high compared with the proportion of deaths in ischemic stroke patients.

From the general analysis, several interlinked signs are evident indicating the likelihood of lethal outcomes in patients with cerebrovascular disease at an early stage: hemorrhagic type of stroke is most likely to be lethal in patients with acute cerebrovascular disease; stroke incidence and mortality are highest in patients aged 60 to 70 years;

stroke with lethal outcomes are more likely in men; regardless of the type of stroke, lethal outcomes are most likely in patients aged 60 to 90 years.

4.2 Mortality day prediction models

All patients were divided into 3 subgroups according to the time of death - up to 1 day, 1 to 3 days, and 4 to 10 days. Early mortality in each subgroup was associated with a number of demographic, clinical, and instrumental-laboratory characteristics based on the interpretation of the results of calculating the significance of predictors of binary classification models by machine learning methods from the Scikit-Learn library². The target classes in training were "mortality rate of 1 day", "mortality rate of 1-3 days", "mortality rate from 4 days". AUC ROC of trained models reached 91% for the method of random forest. The results of interpretation of decision trees and calculation of significance of predictors of built-in methods of random forest coincide that can testify to correctness of calculations. As a result of the decision trees, the following conclusions were drawn regarding the time frame of death:

1. Factors that cause patients to be lethal on the first day: Patient's age over 67 years; male sex; significant volume of brain lesions (more than 1/3 of the middle cerebral artery pool) hemispheric ischemic (or hemorrhagic with impregnation of the ischemic focus) stroke or patients with intracerebral hematoma (both less than 50 ml and 50 to 100 ml) with a breakthrough into the ventricular system of the brain; more important was the combination of ischemic or hemorrhagic lesions with displacement of the medial structures due to perifocal edema; right hemispheric cerebral lesion; severe condition at entry (with severe neurological deficit, up to 23 NIHSS points); unstable systemic hemodynamics, expressed by fluctuations in blood pressure, appearance of tachycardia and tachyarrhythmia, i.e., in the ventricular system.h. with sharp rise (> 200 mm Hg) or sharp decrease (< 65 mm Hg.) systolic and diastolic blood pressure and heart rhythm disorders (tachycardia and tachyarrhythmia); manifestations of decompensated hypersympathicotonia accompanied by hyperthermia and polyuria (densephalic syndrome, irritation of the densephalic region of the brain) and hemoconcentration (hypercoagulation); high degree of comorbidity (presence of significant number of concomitant diseases at the decompensation stage, comorbidity index >6.5).

2. Mortality in the group from 1 to 3 days is caused by the following factors: age over 55 years old; male gender; consciousness impairment not lower than stupor; presence of extensive hemispheric ischemic (more than 1/3 of the middle cerebral artery basin) or large intracerebral hematoma against the background of pronounced brain atrophy, in some cases with hemorrhagic saturation of the ischemic focus; the greatest importance was given to the combination of ischemic or hemorrhagic lesions with displacement of the medial structures due to perifocal edema; lesion of the right hemisphere; instability of system hemodynamics - with indicators of sharp decrease (< 65 mm Hg.st.) of systolic blood pressure, heart rate - with indicators of sharp decrease.st.) systolic blood pressure, heart rhythm disorders (bradiarrhythmia and tachycardia); or with a high degree of comorbidity (presence of a significant number of concomitant

² <https://scikit-learn.org/>

diseases at the decompensation stage, comorbidity index >6.5); vivid manifestations of vegetative regulation decompensation (hypersympathicotonia), accompanied by hyperthermia and polyuria (diencephal syndrome, irritation of the diencephalic region of the brain) and hemoconcentration (hypercoagulation); phenomena of systemic inflammatory reaction and presence of signs of hemoconcentration in blood tests;

3. The largest contribution to the patients' mortality from 4 to 10 days was made by the following factors: age from 30 to 90 years (the largest group of patients aged 60-70 years); female gender; extensive hemispheric ischemic (more than 1/3 of the pool of the middle cerebral artery) in combination with severe hemispheric atrophy, or the presence of intracerebral hematoma (much more often less than 50 ml), a breakthrough into the ventricles of the brain, the most important was the presence of dislocation, a combination of ischemic or hemorrhagic lesions with the displacement of medial structures due to general edema; conscious disturbance (stun, coma) or condition that required sedation (to provide prosthetics for breathing function); unstable systemic hemodynamics - with sharp rise (> 200 mm Hg) or (< 65 mm Hg) of systolic blood pressure; phenomena of moderate hemoconcentration and moderate systemic inflammatory response in blood tests; high degree of comorbidity (presence of a significant number of concomitant diseases at the decompensation stage, comorbidity index > 5). At the same time, it should be noted that in contrast to patients with 1-3 day mortality, in this case the side of the brain lesion did not matter.

4.3 Analysis of mortality factors in subgroups based on P-value assessment

Three groups of patients were compared by the terms of mortality (mortality in the first day, mortality from 1 to 3, lethality from 4 to 10 days) among themselves by means of standard t-test (non-parametric criterion Chi) with thirty one parameter. The value $P < 0.0005$ was considered significant. The results of the interpretation of the obtained test are presented in Tables 1-3.

Table 1. Interpretation of the individual features

Feature	<1 day P	1-3 days P	4-10 days P	Interpretation
Age	0,4190	0,7048	0,6351	Age was not specific (significant) for the development of mortality in the groups under consideration, since the most typical for patients of all subgroups was the age from 60 to 90 years,

Gender	0,0003	0,0004	0,0048	Gender showed the significance of differences between all subgroups, with the groups with mortality of 4-10 days dominated by women, and between the subgroups of mortality up to 1 day and mortality of 1-3 days, with a general prevalence of incidence of men among deceased patients, the frequency of occurrence in the subgroups also significantly differed.
period of admission less than 4.5 hours	0,0003	0,0003	0,7530	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the fact of early hospitalization did not affect the earlier fatal outcome. The differences in subgroups 1-3 and 4-10 are significant
period of admission more than 4.5 hours	0,0003	0,0002	0,5291	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the fact of later hospitalization did not affect earlier mortality. The differences in subgroups 1-3 and 4-10 are significant, for lighter patients (with lower comorbidity index or with severe atrophy).
period of admission over 24 hours	0,0002	0,0002	0,6332	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the fact of later hospitalization did not affect earlier mortality. The differences in subgroups 1-3 and 4-10 are significant, for lighter patients (with lower comorbidity index or with severe atrophy).

Table 2. Interpretation of the Clinical features of the acute period

Feature	<1 day P	1-3 days P	4-10 days P	Interpretation
NIHSS	0,6833	0,0057	0,0401	The difference between subgroups of up to 1 day and 1-3 days is significant, i.e. the fact of more expressed. The differences in subgroups of up to 1 day 4-10 days and in subgroups of 1-3 days and 4-10 days were insignificant.
Clinical features of the acute period				
Charlson comorbidity index	0,0007	0,0007	1,0000	The indicator had the significance of differences when comparing the frequency of mortality in the group 4-10 days with the groups of mortality up to 1 day and mortality 1-3 days, because high values of comorbidity index (>5 and >6.5) played a role in the development of gross and life-compatible disorders in the first 3 days of stroke, which predetermined the lethal outcome.

Consciousness disruption to coma on admission (any quantitative disturbances of consciousness) as well as respiratory failure	0,0001	0,000004	0,3451	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the presence of a coma did not affect earlier mortality. The differences in subgroups 1-3 and 4-10 are significant, for patients entering a coma the treatment program provides for immediate prosthetics of vital functions.
Patients admitted in stupor and remaining conscious for up to 3 days.	0,0002	0,0001	0,3451	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the presence of a coma did not affect earlier mortality. Differences in subgroups 1-3 and 4-10 are significant.
Patients in need of sedation up to 3 days on admission, as well as medical ventilation	< 0,0001	< 0,0001	1,0000	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the necessity to apply sedation is always an unfavorable sign accompanying hypersympathetic reactions, development of diencephalic syndrome or psychomotor excitation, which are associated with high probability of mortality. The differences in subgroups 1-3 and 4-10 are significant, since the use of sedation is aimed at "smoothing" the manifestation of hypersympathicotonia
Instability of system hemodynamics (including bradycardia)	<0,0001	<0,0001	1,0000	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. the presence of instability of system hemodynamics is always a factor associated with early mortality. The differences in subgroups of 1-3 days and 4-10 days are significant, so the appearance of bradycardia and instability of blood pressure are signs of dislocation (contraction) of the central structures of the brain (middle brain, trunk, mediobasal sections of the temporal lobe) - mortality is earlier in these manifestations.
Presence of systemic inflammatory reaction (SIR) or hemoconcentration	0,0003	0,0002	0,5261	The difference between subgroups of up to 1 day and 1-3 days is insignificant. The differences in the subgroups of 1-3 days and 4-10 days are significant, since the development of SIR takes some time, these reactions are signs of a complicated course often accompanied by clinically proven pneumonia, urinary tract infection and/or polyorganic failure syndrome.

Table 3. Interpretation of Morphology

Hemorrhagic stroke	0,0001	0,000016	0,0772	The indicator predetermined a large share of lethal outcomes in the first 3 days, and significantly affected the death in the period up to 1 day and from 1 to 3 days. The indicator was significant when comparing the subgroup 4-10 mortality with the other two.
--------------------	--------	----------	--------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Affection side of hemispheric stroke (as well as aphasia)	0,0003	0,0002	0,7245	The influence of the defeat side on the probability of a lethal outcome is equally significant only when comparing subgroups of 1-3 days and 4-10 days, these differences are more significant for the left hemisphere; differences in the frequency of mortality in subgroups of up to 1 day and 1-3 days and in the left and right hemispheres are insignificant.
Edema	0,0002	0,0002	1,0000	the difference between the subgroups of up to 1 day and 1-3 days is insignificant, i.e. the fact of edema affected earlier mortality. the differences in subgroups 1-3 and 4-10 are significant due to the fact that edema developed later as a factor affecting mortality or did not determine the mortality (e.g. in patients with severe atrophy, small foci).
Dislocation	0,0003	0,0003	1,0000	the difference between the subgroups of up to 1 day and 1-3 days is insignificant, i.e. the fact of edema influenced earlier mortality. the differences in subgroups 1-3 and 4-10 are significant due to the fact that the brain substance dislocation developed later as a factor influencing mortality or also did not determine the mortality (e.g. in patients with severe atrophy, small focus, cortical-subcortical focus, without affecting the central structures of the brain).
Hematoma volume < 50 ml	0,0003	<0,0001	0,000013	The difference between the subgroups is insignificant, i.e. in each case, the fact of the hematoma of a small volume did not in itself cause death, did not have a direct impact on mortality.
Hemorrhagic transformation	0,0003	0,0003	0,3307	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. in each case the fact of hemorrhagic impregnation of the zone of brain matter ischemia affected mortality. In subgroups of 1-3 days and 4-10 days this difference is significant due to availability of reserve spaces due to brain atrophy and less probability of dislocation of brain substance.
Amount of ischemia > 1/3 of the middle cerebral artery (MCA)	0,0002	0,0002	0,4306	The difference between subgroups of up to 1 day and 1-3 days is insignificant, i.e. in each case the fact of extensive ischemic lesion had an impact on mortality in earlier periods. In subgroups of 1-3 days and 4-10 days this difference is significant due to availability of reserve spaces in connection with brain atrophy and less probability of threatening dislocation (constriction) of brain substance even in presence of a large focal point of ischemia and consequently edema and tissue swelling.

Expressed atrophic changes in brain matter	0,0003	0,0003	0,6586	the difference between subgroups of up to 1 day and 1-3 days is insignificant, In the subgroups of 1-3 days and 4-10 days this difference is significant as the availability of reserve spaces due to brain atrophy reduces the probability of dislocation of brain matter even in the presence of a large focal point of ischemia or hemorrhage (large hematoma).
--------------------------------------------	--------	--------	--------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The following calculation results have been obtained P-value using method 2: for the groups 1 and 2 P-Value = 0.0052; for the groups 2 and 3 P-Value = 0.0042; from the groups 1 and 3 P-Value = 0. On the basis of the analysis calculations it is possible to draw a conclusion about a significant difference between the 1st and the 3rd group, where group 1: ICH (Intracerebral hemorrhage) +PVH (periventricular hyperintensity), IS (ischemic stroke) +ICH (intracranial hemorrhage); group 2: IS+Bilat atr, IS-foc 16-25 hu; group 3: IS-1/3<16HU, Sub Tent ICV (Intracerebroventricular).

5 Conclusion and future work

A detailed study of electronic medical records data and combinations of clinical and laboratory characteristics of patients made it possible to reveal dependencies and develop descriptive models between the degree of lesion and the time of intra-hospital lethality of patients. Further, based on a large array of correlated data, models were developed to identify major favorable and unfavorable patterns of early mortality of patients for control and correction of the treatment plan. Decision-making models for predicting the outcome and duration of treatment of stroke patients have been developed using systems analysis, statistical analysis, mathematical modeling and machine learning methods. As a result, clinical and morphological predictors of early hospital stroke mortality have been identified. Similar models can also be used to validate existing scales, to study the causes of mortality at the emergency stages and to develop clinical guidelines, including for the prevention, diagnosis and treatment of stroke.

As a result of this study, descriptive and prognostic models of mortality in stroke patients have been developed. The significance of predictors was ranked using statistical and machine learning methods. Clinical interpretation of the obtained results was made in the form of clear conclusions that can be used in the organization of continuity care for acute stroke patients, as well as the calculation of personal risks. Provided that all standards of specialized medical care for patients with stroke are complied with, first of all, monitoring and intra-hospital logistics, completeness of the diagnostic scope, it is possible to make a prognostic assessment to identify predictors of early hospital lethality. A number of clinical, pathomorphological and instrumental parameters may indicate a high probability of early lethality, namely: Charlson comorbidity index with a value greater than 3.0; six subtypes of stroke; for 3 subtypes, combination with an extended intracellular CMA clot, with an age greater than 64 years and Ind. Ch-4.5 b. For 1 and 2 subtypes the severity of lesion volume and presence of dislocation complications determine the high risk of mortality. For 4 subtypes the greatest risk is associated

with the combination of an acute focus in the deep parts of the temporal lobe with moderate perifocal ischemic oedema, compression of medial structures, with the age over 64 years old and high and Ind. Ch-4.5 b. For subtype 6, a significant contribution is made by global (diffuse atrophy or the presence of a fresh acute focus in the deep regions of the temporal lobe on the side of the opposite marked atrophy (including post-stroke).

Aknowledgements

This work was financially supported by the Government of the Russian Federation through the ITMO fellowship and professorship program. This work is financially supported by National Center for Cognitive Research of ITMO University.

References

1. Johnson C.O. et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016 // *Lancet Neurol.* 2019.
2. Jdanov D. et al. Human Mortality Database // *Encyclopedia of Gerontology and Population Aging.* 2019.
3. Global burden of stroke and risk factors in 188 countries, during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013 // Elsevier.
4. Feigin V. et al. Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010 // Elsevier.
5. Lee M. et al. Impact of microalbuminuria on incident stroke: A meta-analysis // *Stroke.* 2010. Vol. 41, № 11. P. 2625–2631.
6. Wang Y. et al. The China National Stroke Registry for patients with acute cerebrovascular events: design, rationale, and baseline patient characteristics // *Int. J. Stroke.* 2011. Vol. 6, № 4. P. 355–361.
7. Sun S. et al. GWTG Risk Model for All Stroke Types Predicts In-Hospital and 3-Month Mortality in Chinese Patients with Acute Stroke // Elsevier.
8. Fekadu G., Chelkeba L., Kebede A. Burden, clinical outcomes and predictors of time to in hospital mortality among adult patients admitted to stroke unit of Jimma university medical center: a prospective cohort study // *BMC Neurol.* Springer Science and Business Media LLC, 2019. Vol. 19, № 1.
9. Chen H. et al. Analysis on geographic variations in hospital deaths and endovascular therapy in ischaemic stroke patients: an observational cross-sectional study in China // *bmjopen.bmj.com.*
10. Alemayehu Gebreyohannes E. et al. In-Hospital Mortality among Ischemic Stroke Patients in Gondar University Hospital: A Retrospective Cohort Study // *Stroke Res. Treat.* 2019.
11. Heuschmann P.U. et al. Predictors of In-Hospital Mortality and Attributable Risks of Death After Ischemic Stroke The German Stroke Registers Study

- Group.
12. Koennecke H. et al. Factors influencing in-hospital mortality and morbidity in patients treated on a stroke unit // *AAN Enterp.*
 13. Minnerup J. et al. Explaining the decrease of in-hospital mortality from ischemic stroke // *PLoS One. Public Library of Science*, 2015. Vol. 10, № 7. P. 1DUUMY.
 14. Smith E.E. et al. Risk score for in-hospital ischemic stroke mortality derived and validated within the get with the guidelines-stroke program // *Circulation*. 2010. Vol. 122, № 15. P. 1496–1504.
 15. Cadilhac D.A. et al. Risk-Adjusted hospital mortality rates for stroke: Evidence from the Australian Stroke Clinical Registry (AuSCR) // *Med. J. Aust. Australasian Medical Publishing Co. Ltd*, 2017. Vol. 206, № 8. P. 345–350.
 16. Steiner M.M., Brainin M. The quality of acute stroke units on a nation-wide level: The Austrian Stroke Registry for acute stroke units // *European Journal of Neurology*. 2003. Vol. 10, № 4. P. 353–360.
 17. Stepanova M. et al. Recent trends in inpatient mortality and resource utilization for patients with stroke in the United States: 2005-2009 // *Elsevier*.
 18. Keller K. et al. Impact of atrial fibrillation on in-hospital mortality of ischemic stroke patients and identification of promoting factors of atrial thrombi—Results from the German // *ncbi.nlm.nih.gov*.
 19. Ganesh A. et al. Integrated systems of stroke care and reduction in 30-day mortality: a retrospective analysis // *AAN Enterp.*
 20. Counsell C. et al. Predicting outcome after acute and subacute stroke: Development and validation of new prognostic models // *Stroke*. 2002. Vol. 33, № 4. P. 1041–1047.
 21. Gattringer T. et al. Predicting Early Mortality of Acute Ischemic Stroke: Score-Based Approach // *Stroke*. – 2019. – T. 50. – №. 2. – C. 349-356.
 22. Ho W.M. et al. Prediction of in-hospital stroke mortality in critical care unit // *Springerplus. SpringerOpen*, 2016. Vol. 5, № 1.
 23. Lee J. et al. Derivation and validation of in-hospital mortality prediction models in ischaemic stroke patients using administrative data // *karger.com*.
 24. Weimar C. et al. Age and National Institutes of Health Stroke Scale Score Within 6 Hours after Onset Are Accurate Predictors of Outcome after Cerebral Ischemia: Development and External Validation of Prognostic Models // *Stroke*. 2004. Vol. 35, № 1. P. 158–162.
 25. Baptista M. et al. Prediction of in-hospital mortality after first-ever stroke: the Lausanne Stroke Registry // *Elsevier*.
 26. Wang Y. et al. A prognostic index for 30-day mortality after stroke // *Elsevier*.
 27. Tanne D. et al. Trends in management and outcome of hospitalized patients with acute stroke and transient ischemic attack: The national acute stroke israeli (nasis) registry // *Stroke*. 2012. Vol. 43, № 8. P. 2136–2141.